

CORRESPONDENCE

Non-smoking wives of heavy smokers have a higher risk of lung cancer

H R Kornegay and M A Kastenbaum, PHD; N Mantel; J E Harris, MD, and W H DuMouchel, PHD; Eleanor J Macdonald, MD; T Hirayama, MD..... 914

Statistics in question

M F Grayson, MRCP..... 917

Caution on preventing neural tube defects

V C Talwalkar, MD..... 917

The right to live and the right to die

R B Zachary, FRCS..... 918

Microcomputers in antenatal care

J R Davies, FFARCS..... 918

Diving is dangerous for diabetics

J D King, MB..... 918

Dealing with epileptics

Rachel Tavriger..... 918

Management of asthma in the child aged under 6 years

I Blumenthal, MRCP..... 918

Minor orthopaedic problems in children

A N Conner, FRCS, and J E Robb, FRCS..... 919

Reversible renal damage due to glue sniffing

Mary D King, MB..... 919

Audit in renal failure

H J Goldsmith, FRCP, and others..... 919

Vitamin A toxicity and hypercalcaemia in chronic renal failure

M O'Fearghail, MB, and others..... 919

Treatment of acute mountain sickness

P McDonald, FRCS..... 919

Proposals for a trial marriage between primary and secondary health care in one or two districts in inner London

C J Dickinson, FRCP..... 920

Junior posts and expansion of the consultant grade

A W Banks, FRCOG..... 920

Changes in the MRCP examination

M Ross, MRCP, and D M F Gibb, MRCP... 920

Will doctors miss out again?

R C L Feneley, FRCS..... 920

Doctors' contracts: an urgent case for legislation

J L Taylor, MRCP; F E Weale, FRCS..... 921

Consultant contracts

B A Kamdar, FRCS..... 921

New technique of drug promotion?

J Iqbal, MRCP..... 921

Corrections: Hypokalaemia due to salbutamol overdosage (Corea *et al*); Dealing with epileptics (Beran and Sutton)...

Non-smoking wives of heavy smokers have a higher risk of lung cancer

The United States Tobacco Institute issued a press release in June disputing the conclusions of Dr T Hirayama's paper "Non-smoking wives have a higher risk of lung cancer," published in the *BMJ* of 17 January 1981. This received wide publicity. We have received many letters and other documents on the subject and have therefore decided to reopen our correspondence. The following letters were sent to Dr Hirayama, who replies at the end.

Although two of the communications were not written as letters to the *BMJ* they are an essential part of the story and we have taken the exceptional step of printing them as they stand.—Ed, *BMJ*.

SIR,—Dr Takeshi Hirayama (17 January, p 183) claimed a higher risk of lung cancer among non-smoking wives of smokers than in non-smoking wives of non-smokers in his Japanese study population. Other scientists in Europe and the United States have since questioned the study (28 February, p 733; 21 March, p 985; 4 April, p 1156; 25 April, p 1393). In a larger population Garfinkel¹ has found no such higher risk among United States women.

Meanwhile, several US experts had found an apparent statistical error in the Japanese calculations—raising serious questions about the study. We regard this discovery as very grave, particularly because of the effect in the United States, where popular media and legislative bodies have used the presumption of danger expressed in the Japanese study as a rationale for regulations designed to restrict smoking in public areas.

Early in March 1981 we submitted many questions to the author of the study—questions which addressed major scientific concerns about the study. For reasons not known to us, the author chose not to reply to these questions. Among the most important questions raised in this correspondence were these:

(1) "You report that mortality rates are 'age-occupation standardised annual mortality rates.' I have not been able to reproduce these numbers because the ages and occupations of

your subjects are not available to me, nor do I know the population against which your data were standardised. Could you please furnish this information so that I may reproduce the reported rates?"

(2) "One of your groups of husbands included both non-smokers and occasional smokers. How is an occasional smoker defined? Why were these two classes combined? Another group of husbands consisted of ex-smokers and smokers of 19 or fewer cigarettes daily. How many ex-smokers were included in this group? Why were these two classes combined?"

(3) "Did you measure the times, if any, that the non-smoking wives were present when their husbands smoked? If yes, how was this done?"

In light of these questions and others raised by members of the scientific and medical communities, we believe the claims in the Japanese study to be unsubstantiated.

HORACE R KORNEGAY

Chairman

MARVIN A KASTENBAUM

Director of statistics

Tobacco Institute,
Washington DC 20006

¹ Garfinkel L. *J Nat Cancer Inst* 1981;**66**:1061-6.

To Dr Marvin A Kastenbaum, Tobacco Institute

This paper by Dr Takeshi Hirayama (17 January, p 183) reports certain results of a major prospective investigation on the effects of cigarette smoking, the novel feature emphasised in the paper being on the implications relating to indirect or passive smoking and lung cancer.

As part of the report, the author gives standardised lung cancer mortality rates for women subdivided by the smoking habits of their husbands or by their husbands' ages or both, presumably at the initiation of the study. By standardised mortality rates the author would have meant age-standardised mortality rates. To have properly obtained such standardised rates the author would have had to consider each individual year of life to which a

woman survived, and presumably he did so. The data covered the period 1966-79.

The statistical analysis of the resulting data would be rather complex, though readily handled by procedures which I have published, and some short cuts are possible. Because of the rapid rise of cancer rates with age, I would stratify the non-smoking married women into comparatively narrow initial age groupings, say two years in width. Each age grouping would be followed year by year so as to contribute information on persons at risk in each of the three husbands' smoking categories (non-smoker, light smoker or ex-smoker, heavy smoker—that is, ≥ 20 cigarettes per day; since only individuals 40 and over were initially recruited, we can ignore any change in husband's status), and also on the number of women among them dying of cancer each year. Since the women are homogeneous on age initially, they will continue to be so over the entire period.

Other factors would be readily incorporated into this analysis. Dr Hirayama emphasises in particular as stratifying factors the husband's initial age, 40-59 years versus ≥ 60 years, and the husband's broad occupational grouping, agricultural versus non-agricultural. Either, neither, or both of these factors can be incorporated into the analysis. Actually, there would be 12 possible analyses: adjusting for neither A nor B; adjusting only for A; adjusting only for B; adjusting for both A and B; examining separately in each of two levels of A ignoring B; examining separately in each of two levels of B ignoring A; examining separately in each of the four A \times B combinations.

To some extent Dr Hirayama has done something much akin to this. He reports in tabular form summary data which would correspond to six of the 12 possible analyses, while giving in the text the data which would correspond to yet a seventh. For statistical methodology he notes my extended χ^2 procedure, which would suggest that he is treating the wives of light smokers or ex-smokers as being at an exposure level midway between non-smokers and heavy smokers. But what is surprising is that he makes no allusion to what age groupings he used in making his statistical analyses or how he took into account the passage of time in making them. Perhaps they mattered in calculating standardised rates, but not when he performed statistical analyses. (But I have my doubts even here—in one part of table I Dr Hirayama refers to occupation-standardised mortality, in another to age-standardised mortality. But the occupations characterise the husbands, not the wives, and so also might the ages.)

Anyway, let's have a look at some of the summary data given. Collectively, for wives of non-smokers we have 32 lung cancers among 21 895 women versus 86/44 184 for ex-smokers or light smokers versus 56/25 461 for heavy smokers. That would yield a single df χ^2 by the method Dr Hirayama apparently used of 3.31, or a χ value of 1.82, not significant, two tailed, at the 5% level. Dr Hirayama, however, reports a χ value of 3.30, with a two-tailed probability level beyond 0.1%. The question then is whether he has conducted a more refined analysis, about which he is giving us no clues, or he has mistakenly interpreted his χ^2 value as a χ value.

As an example of a particularly striking association, Dr Hirayama concentrates on the results for non-smoking wives of agricultural workers aged 40-59. The corresponding data for this group are 3/5 999 versus 20/12 753 versus 16/7 150, which would yield a χ^2 of 6.45, a χ value of 2.54. Dr Hirayama reports a surprisingly close χ value of 2.60, with an associated significance level of 0.94%. As an extreme outcome of several possible analyses available to him this is not really all that significant. But if the difference between 2.54 and 2.60 reflects only rounding errors in calculations rather than distinctly different analyses, this would make it more likely that Dr Hirayama's 3.30 value is a χ^2 rather than a χ , which, as such, does not have the extreme statistical significance which he has attached to it.

Much more careful analysis of the data would be needed before it can be claimed that a passive effect of smoking has been clearly established.

NATHAN MANTEL

Department of Statistics,
Biostatistics Center,
George Washington University,
Bethesda, Maryland 20014,
USA

To Dr Takeshi Hirayama

In a recent press release the Tobacco Institute in the United States challenged the statistical validity of your finding that non-smoking wives of heavy smokers have a higher risk of lung cancer. The central point raised by the Tobacco Institute was that you erred fatally in your calculation of a critical test statistic and therefore that your claim of a high level of statistical significance is wrong. The Tobacco Institute's assertions relied heavily on a speculation contained in a memorandum on your study by Dr Nathan Mantel. The purpose of this letter is to inform you of the details of Dr Mantel's analysis, and in particular the origin of the alleged arithmetical error. Your results, at least those published, do indeed display a high level of statistical significance. If there is an error of inference in your study, it is far less superficial than the alleged arithmetical mistake publicised by the Tobacco Institute.

The essential data are given in the attached table, which we derived from tables I and II of your paper in the *BMJ* (17 January, p 183). The population of non-smoking wives at the beginning of the prospective study and the number of lung cancer deaths during 14 years of observation are broken down according to the age, occupation, and smoking habits of their husbands. Also provided are the corresponding 14-year lung cancer death rates for each cell. At the bottom of the table data for all age and occupational categories of husbands are combined.

In his memorandum to the Tobacco Institute, Dr Mantel calculated a χ^2 test statistic with one degree of freedom from the combined data, and obtained the result $\chi^2 = 3.31$, which would yield a two-sided $p = 0.07$. In performing this test he used the scoring 0, 1, 2 for the three smoking categories. This value of χ^2 looked suspiciously close to the value of $\chi = 3.299$ reported on page 183 of your paper. On this basis, Dr Mantel speculated that

you may have failed to take a simple square root in your calculation.

If, however, a χ^2 test statistic with one degree of freedom is calculated from the disaggregated data in the table, then the result is $\chi^2 = 8.09$, which yields a highly significant two-sided $p = 0.004$. In making this calculation we used formulas (2) through (4) on page 694 of Dr Mantel's original paper,¹ which shows how to combine χ^2 statistics from the four independent contingency tables. The reason for the increase in statistical significance is that the dose-response relation between the husband's smoking habit and the non-smoking wife's lung cancer rate, which is apparent from the combined data, becomes even more prominent when the data are broken down by age and occupational categories of the husband. That is exactly what Dr Mantel's extension χ^2 was designed to test.

We reach the same conclusion of a high level of statistical significance when the disaggregated data are fitted to a number of plausible statistical models. For example, if P_{ij} denotes the 14-year lung cancer death rate of a non-smoking woman whose husband belongs to age-occupational category $i = 1, 2, 3, 4$ and smoking category $j = 1, 2, 3$, then the traditional relative risk model $P_{ij} = a_i \cdot r_j$ yields a good statistical fit with an equally strong level of significance. (The 6-degree-of-freedom χ^2 goodness-of-fit statistic is 5.3, with $p = 0.5$.) For this model, the estimated relative risk of lung cancer for the intermediate category of smoking husbands is 1.42 (95% confidence interval 0.95 to 2.13) by maximum-likelihood methods. The estimated relative risk of lung cancer for the heavy-smoking husbands is 1.87 (95% confidence interval 1.21 to 2.91). These estimates are consistent with the relative risks reported in your paper.

In fairness and respect to Dr Mantel, we wish to put out that his speculation of an arithmetical error was only a minor point in his memorandum. The greater part of Dr Mantel's criticisms was devoted to ambiguities in your method of presenting the results. Thus it is not entirely clear to us whether you based your statistical calculations on more finely disaggregated age and occupational groups than those reported in the paper and reproduced by us here. It is unclear whether you used the number of person-years at risk or the number of persons at risk in your calculations of statistical significance. It is also not clear to us whether you standardised on the ages of the wives themselves. Such calculations, if not already performed, would certainly make the analysis more conclusive. But we see no obvious reason why these additional analyses should produce substantially different results.

The Tobacco Institute's claim that your study is invalidated by a trivial arithmetical error has become an international news story. We recognise that you, as author of the paper

under challenge, are in the most appropriate position to respond publicly. As your colleagues in the scientific community, we urge you to uphold the validity of your paper.

JEFFREY E HARRIS
WILLIAM H DU MOUCHEL

Department of Economics,
Massachusetts Institute of
Technology,
Cambridge, Mass 02139,
USA

¹ Mantel N, Halperin M. *Journal of the American Statistical Association* 1963;58:611-27.

SIR,—The far-reaching implications of the paper on passive smoking (17 January, p 183) make it imperative to determine whether the data on which it is based are adequate for the purpose.

In none of his several reports on this population does Hirayama explain how he selected his study population. There are 46 prefectures in Japan. Within these prefectures there are 832 health stations. In the six prefectures from which the 69 health stations were selected there are 173 health stations. How were the 69 selected? In one report it was stated that they were the medically adequate ones, which would introduce certain biases. The six prefectures included in the study contain most of the heavy industries in Japan and also border the sea. Shipbuilding, for example, in which Japan leads in world production, is associated in the United States with asbestos exposure. Since shipbuilding is a major industry in one of the prefectures in the study, were the health stations selected for the study for residents in the shipbuilding areas? Industry in Japan supplies medical care to workers and their families as one of the perquisites of employment. Were the families covered by industry omitted from the study? If so, the healthiest segment of the population was omitted.

The figure of 91-99% of the population of the health stations is given to demonstrate the population basis of these data. To what population is Hirayama referring? Is the census in Japan taken of health stations—that is, of individuals eligible for care in health stations? Is not attendance at the health stations the option of the individual? In the *Japanese Almanac* for 1972, the proportion of females visiting the health stations ranges from 2.5% in Miyaga to 14.5% in Osaka. How can this be population based? And if it is population based why were the rates figured by the person-year method rather than on the actual population?

Personal interview surveys usually elicit more complete response than mailed questionnaires or hospital record searches. In this study, on the pivotal question of occupation 44 357 (31.04%)

Lung cancer deaths in non-smoking Japanese wives during 14 years of observation: data based on tables I and II of Hirayama's paper*

Husband's smoking habit:	Non-smoker	Ex-smoker or 1-19 cigarettes/day	≥20 cigarettes/day
<i>Disaggregated data</i>			
Husband aged 40-59 years and working in agriculture:			
Population of wives at risk	5999	12 753	7150
No of deaths from lung cancer	3	20	16
Lung cancer deaths per 10 000	5.0	15.7	22.4
Husband aged 40-59 years and working elsewhere:			
Population of wives at risk	8021	17 923	13 434
No of deaths from lung cancer	8	20	20
Lung cancer deaths per 10 000	10.0	11.2	14.9
Husband aged 60 years or more and working in agriculture:			
Population of wives at risk	4407	7291	2241
No of deaths from lung cancer	14	32	8
Lung cancer deaths per 10 000	31.8	43.9	35.7
Husband aged 60 years or more and working elsewhere:			
Population of wives at risk	3468	6217	2636
No of deaths from lung cancer	7	14	12
Lung cancer deaths per 10 000	20.2	22.5	45.5
<i>Combined data</i>			
Husbands of all ages and occupations:			
Population of wives at risk	21 895	44 184	25 461
No of deaths from lung cancer	32	86	56
Lung cancer deaths per 10 000	14.6	19.5	22.0

* *BMJ* 1981;282:183-5.